Physiological Chemistry and Physics and Medical NMR, 18 (1986)

THE ROLE OF INDUCTIVE EFFECT IN THE DETERMINATION OF PROTEIN STRUCTURE

GILBERT N. LING
Department of Molecular Biology, Pennsylvania Hospital, Eighth and Spruce Streets, Philadelphia, PA. 19107.

• ***The inductive theory, formally introduced by G. N. Lewis has thus far found its major applications in interpreting and predicting equilibrium and kinetic properties of small organic molecules. Evidence is presented demonstrating that the inductive effect can also help to explain the determination of protein structure by its amino acid sequence. Suggestions are also made that the inductive effect plays a significant role in protein conformation changes brought about by ligand binding.***

Renaturation experiments showed that the secondary and high-order structure of a protein are determined by its amino acid sequence (Anfinsen, 1962, 1967). By examining proteins of known amino acid sequences and helical contents, investigators discovered rules that govern the formation of a-helical structure and non-helical structures. Certain amino acid residues occur frequently in the a-helical segments while others are excluded from them (Szent Györgyi & Cohen, 1957; Davies, 1964; Guzzo, 1965; Edmundson, 1965; Havsteen, 1966; Prothero, 1966; Periti et al., 1967; Low et al., 1968; Goldsak, 1969). An important deduction of Scheraga and his coworkers (Kotelchuck & Scheraga, 1969; Lewis & Scheraga, 1971) is the principle of predominance of short-range interaction which Scheraga described in these words, "the conformation of an amino acid residue in a polypeptide or protein is determined in very large measure (though not exclusively) by the short-range interactions between a side chain and the atoms of the backbone of the same amino acid residue . . ." (Scheraga, 1974, p. 1; see also Finkelstein & Ptitsyn, 1971).

With x-ray crystallographic data of more proteins available, quantitative parameters were obtained describing the different potentials of each amino acid residue toward assuming an a-helical, β-pleated sheet, turns or coil structure in a polypeptide or protein. With these quantitative parameters on hand, the secondary structures of a protein have become predictable from its amino acid sequence (Chou & Fasman, 1974, 1978; Tanaka & Scheraga, 1976; Garnier et al., 1978).

As time went by, it has become clear that the secondary structure of a specific segment of polypeptide is also influenced by the neighboring amino acid residues close by (intermediate range interactions) or far removed (Wu & Kabat, 1971; Robson & Pain, 1972; Ponnuswamy et al., 1973; Lim, 1974; Nagano, 1974; Robson & Suzuki, 1976). As pointed out by Lim (1974), secondary structures must have dimensions limited by those of the globular molecule as a whole. Further, they must contribute to minimize the total free energy by the formation of a compact hydrophobic core protected from contact with the aqueous environment with a surface of hydrophilic groups (Kendrew et al., 1960). Electrostatic dipole-dipole interactions between α-helices (Wada, 1976; Sheridan et al., 1982; Rogers & Sternberg, 1984) also influence

secondary structure choices. All these considerations point to the role of long-range forces in determining the stability of specific secondary structures.

Using methods introduced in information theory, Robson and his colleagues made some important discoveries. They demonstrated that the influence of an amino acid residue (j) in a protein is not limited to the conformation of the peptide belonging to the jth amino acid residue but it also influences the conformations of amino acid residues at j−m and j+m positions along the polypeptide chains where m varies from 0 to 8. (Residues labelled j−m are on the N-terminal side of the polypeptide chain; j+m on the C-terminal side.) When $m \neq 0$, the component of the information measured is referred to as "directional information" (Robson & Pain, 1972). Robson and Pain (1974) found that in globular proteins they studied, the conformations of approximately half of the residues were determined by their side chains (short-range interaction) whereas in the remaining half, this conformational preference dictated by the side chain was overridden by interactions between each residue and other residues in the protein (intermediate and long-range interactions).

SEVERAL THEORETICAL PREDICTIONS BASED ON AN EXTENSION OF THE INDUCTIVE THEORY TO PROTEIN STRUCTURE

That various empirically derived formulas can enable one to predict secondary structure with an accuracy as high as 60% or even 80% is impressive. Undoubtedly with additional information from further tertiary structure analyses (Lim, 1974; Taylor & Thornton, 1983; Welinder et al., 1984) still more refined rules can be obtained to generate even more accurate predictions of protein structures. Yet the ability of a specific (jth) amino acid residue to form hydrophobic bonds, salt-linkages, H-bonds, etc., requires the presence of other suitable "partners" at the proper location — a presence which is not under the control of the jth residue and rightly belongs to "intermediate range interactions". In contrast, the short-range interaction of a specific amino acid residue is entirely independent and describable by quantitative parameters true to that amino acid residue in whatever protein it is found. For this reason, it seems that another more independent mechanism is needed in order to explain the dominant short-range component of the potentials for a-helical and other conformations. It is with the hope of providing a candidate for this mechanism, that I reexamined and extended a suggestion I made some years ago. In brief, an inductive effect might play a role in the control by a side chain of the choice of conformation of its own peptide amide group as well as the conformations of immediately neighboring peptide amide groups (Ling, 1964, 1984).

The suggestion I made involves four postulations:

Postulate 1. The formation of an α-helical or β-structure of a specific peptide amide group rather than a random coil is primarily a consequence of the more favorable free energy in the formation of the H bonds between a specific pair of peptide amides than in the formation of H bonds between these specific peptide amide groups and water molecules.

Suggestive Evidence: At room temperature many proteins conserve their a-helical structures in the presence of 55 M water (while others do not); yet they may denature and assume the random coil conformation on exposure to urea five times less concentrated. Kauzmann (1954) long ago suggested that it is the stronger H-bonding power of urea than that of water which underlies the much greater effectiveness of urea than water in breaking up native secondary structures. The greater

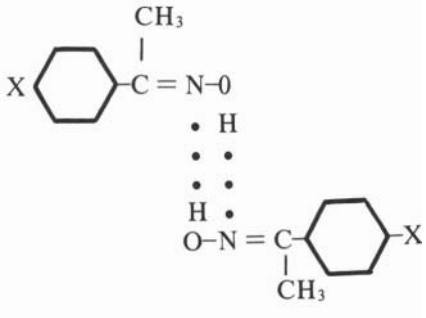
effectiveness of urea than water in denaturing proteins also points out that in these proteins the H-bonding strength follows the rank order: peptide-urea > peptide-peptide > peptide-H_2O.

Postulate 2. The strength of the peptide-amide to peptide-amide H bonds varies; it depends primarily on the basicity (i.e., electrondonating power) of the amide carbonyl groups and to a much less extent on the acidity (i.e., electron-accepting power) of the amide NH groups.

Suggestive Evidence: Infrared spectrum studies showed that the C = 0 bond of the amide group is very polar whereas the NH bond is not (Cannon, 1955), and that the C = 0 group is a strong electron donor whereas the NH is a weak electron acceptor (Mizushima, 1954).

Postulate 3. The basicity of the peptide amide C = 0 groups is determined primarily by the relative electron-donating power of the contiguous side chain and ligand associated with the side chain, e.g., H^+ or Na^+ bound to a β-carboxyl group.

Suggestive Evidence: Dimerization of acetophenone oximes involves the formation of a pair of H bonds not unlike those formed between pairs of peptide amides in a-helical or β-pleated sheet conformations:



The free energy of dimerization was found to be linearly related to the Hammett's inductive constant of substituent X at the para position of the aromatic rings. Electron-donating substituents (e.g., CH_3) increase the free energy of the H bonds; electron-withdrawing para substituents (e.g., NO_2) decrease it. Reiser (1959) concluded that it was the electron density (i.e., basicity) at the electrondonating nitrogen atom that determined the H-bonding strength. In H bonds formed between peptide amides, the role of N and O atoms are reversed: it is the carbonyl oxygen of the peptide amide that is electron-donating and thus equivalent to the N atom in acetophenone oxime.

Postulate 4. The electron-donating power of a specific side chain (and ligand associated with the side chain) does not influence only the amide group of this amino acid residue but may affect peptide amide groups (and some side chains) of the immediately neighboring amino acid residues.

Suggestive Evidence: The effectiveness of the polypeptide chain to transmit the inductive effect is indicated by its partially resonating structure (Pauling, 1960) and by the demonstration (in water and 1 N NaCl) of continuing, monotonic but diminishing changes of the acid dissociation constants of the a-amino groups and a-carboxyl groups at the two ends of the molecules separated by an increasing number of peptide linkages in the series: glycine, diglycine, triglycine, tetraglycine, etc. (Stiasny & Scotti, 1930; Czarnetzky & Schmidt, 1931; Greenstein & Winitz, 1961; Ling, 1962).

The following predictions can be deduced from the above extension of the induction theory first formally proposed by G. N. Lewis (1923) and further developed by many others (Hammett, 1935, 1970; Branch & Calvin, 1941; Dewar, 1949; Smith et al., 1951; Ingold, 1953; Taft, 1960; Chiang & Tai, 1963; Chapman & Shorter, 1972):

(1) The electron-donating power of the

Amino Acid Type	Carboxylic Acid Analogues	Acid Dissociation Constants A (pK_a)	Acid Dissociation Constants B ($\overline{pK}_a$)	sources
Asp	malonic acid	5.69 (3.43*)	4.56	(1)
Glu	succinic acid	5.61 (4.76*)	5.19	(1)
Ile	a-methyl-n-butyric acid	4.78	4.78	(2)
Val	isobutyric acid	4.82	4.82	(1)
Leu	isovaleric acid	4.77	4.77	(1)
Ala	acetic acid	4.75	4.75	(1)
Try	indoleacetic acid	4.75	4.75	(3)
Gln	succinamic acid	4.60	4.60	(4)
Met	methylthiopropionic acid	4.50	4.50	(7)
Tyr	p-hydroxylphenylacetic acid	4.28	4.28	(6)
Phe	phenylacetic acid	4.25	4.25	(3)
Arg	guanidylbutyric acid	4.35 (4.8**)	4.58	(7)
Lys	6-amino-n-valeric acid	4.2 (5.2**)	4.7	(5)
Thr	lactic acid	3.86	3.86	(3)
Ser	glycolic acid	3.80	3.80	(1)
Gly	formic acid	3.75	3.75	(1)
Cys	thioglycolic acid	3.67	3.67	(3)
Asn	malonamic acid	3.64	3.64	(4)
His	imidazoleacetic acid	3.3 (3.95**)	3.63	(6)

TABLE I.'' ***The acid dissociation constants of the a-carboxylic acid analogues of 19 a-amino acids.*** The analogue of each a-amino acid, $R\text{-}CH(NH_2)COOH$, chosen is its corresponding a-carboxylic acid, RCOOH. Under Column A, the charged "side chains" of the analogues are in the fully ionized, charged form unaccompanied by counterions of any kind (e.g., $COO^-\text{-}CH_2COOH$; $NH_3^+(CH_2)_4COOH$). The acid dissociation constants marked with asterisks (in parentheses) are those of the analogue acids with the side chains in an unionized form ($COOH\text{-}CH_2\text{-}COOH$;

Source of a-helical potential data	Correlation Coefficients with the Acid Dissociation Constants			
	a	b	c	d
Tanaka and Scheraga	+0.59	+0.84	+0.81	+0.75
Chou and Fasman	+0.64	+0.78	+0.81	+0.77
Garnier, Osguthorpe, and Robson	+0.60	+0.76	+0.80	+0.72

TABLE II.[a] Linear correlation *coefficients of* acid *dissociation constants of the* a-carboxylic analogues *of* rhe 19 a-amino acids *against the* α*-helical potentials of* the *different* amino acids. a: pK_a data as given under Column A of Table I. b: pK_a data as given under Column A of Table 1 minus all five charged groups, asp, glu, lys, arg, and his, c: pK_a data as given under Column A of Table I minus only asp and his. d: $\overline{pK}_a$'s as given under Column B of Table I (see text for explanation). a-helical potentials are those originally given as ω_{jh}* (Tanaka & Scheraga, 1976), P_α from 29 proteins (Chou & Fasman, 1978) and directional information for the a-helical conformation at the residue position j in units of centinats (Garnier, et al., 1978). For the sample size of 19, a 1% two-tailed significance corresponds to an r of 0.575 (Snedecor & Cochran, 1980). The present data are, of course, considerably better than that.

different side chains of various amino acid residues should show positive correlation with the respective helix-forming potentials of the amino acid residues.

(2) The electron-donating power of the different side chains of the various amino acid residues should also show positive correlation with their respective β-structure forming potentials though this correlation should be weaker than that with the helix potentials. This difference arises from the conservation of electrons donated by the side chains to within the same polypeptide chain in an a-helical conformation in contrast to its partial dissipation among neighboring chains in β-structures.

(3) The electron-donating power of the different side chains of the various amino acid residues should demonstrate negative correlation with their respective coil-forming potentials.

(4) The electron-donating power of the jth residue should also show positive correlation with a-helix and β-structure forming potential of immediately neighboring peptide bonds.

← $NH_2(CH_2)_4COOH$). For asp and glu, the pK_a's of the "undissociated" a-carboxylic groups are the first dissociation constants of the dicarboxylic acid to which the value 0.6 (= RT ln 2) has been added (see text). The pK_a's of "undissociated" arg, lys, and his marked with double asterisks, were calculated according to the method of Chiang and Tai (1963) from the molecular structures of the analogue a-carboxylic acids (see text). Sources of other data were as follows: (1) Hodgman et al. (1961); (2) Jencks & Kegenstein (1970); (3) Stecher (1968); (4) Kortum et al. (1961); (5) Branch & Calvin (1941); (6) new data obtained by titration (Ling & Ochsenfeld, unpublished); (7) estimate based on pK_a = 3.72 of methylthioacetic acid (Jencks & Kegenstein, 1970). It is known that pK_a values of derivatives of acetic acids (XCH_2COOH) and n-propionic acids (XCH_2CH_2COOH) are linearly related to Taft's inductive constants (σ_I) of the substituents X (Taft & Lewis, 1958; Ling, 1964). From these $\overline{pK}_a$'s vs. σ_I plots, one can estimate the pK_a of XCH_2CH_2COOH from the known pK_a of XCH_2COOH. By using this method the pK_a rnethylthiopropionic acid was obtained based on the known pK_a of methylthioacetic acid (3.72) (Jencks & Kegenstein, 1970). We titrated guanidylpropionic acid, yielding a pK_a of 3.8. By a similar method as described above under (7), we obtained linear plots of pK_a of $XCH_2CH_2CH_2COOH$ against σ_I. Comparing this plot with the XCH_2CH_2COOH vs. σ_I plot one estimated the pK_a of guanidylbutyric acid.

RESULTS

Many superb indices of helix potentials, β-structure potentials, etc. of different amino acids are already available. The initial effort needed to test the hypotheses is to create a list of electron-donating powers of the different amino acid side chains; this was obtained in the following manner: It is well known that the pK_a's of carboxylic acids RCOOH are linearly related to the inductive constants of R groups (Derick, 1911; Branch & Calvin, 1941). That is, a strong electron-donating group like CH_3 in acetic acids (the analogue of alanine) produces a carboxyl group with a high pK_a value (4.75), whereas a weak electron-donating group, H, in formic acid (the analogue of glycine) produces a carboxyl group with a low pK_a value (3.75). Since the strong mutual interaction between the α-amino and a-carboxyl groups of the free amino acids precludes the use of the acid dissociation constants of the a-carboxyl groups or a-amino groups of the free amino acids themselves, I chose to use the pK_a's of a-carboxylic acids with "side chains" corresponding to those of the 19 free a-amino acids making up most proteins (proline, not an a-amino acid, is excluded). Thus the analogous a-carboxylic acid to alanine, CH_3CHNH_2COOH is acetic acid, CH_3COOH; that of tryptophane is indole-acetic acid, etc. These and other analogue acids and their pK_a's are listed in the second and third columns of Table I.

While most of the pK_a values were available in the literature, five had to be determined and/or calculated from existing pK_a's of closely related acids. The details are given in the legend of Table I.

Column a of Table II shows the linear correlation coefficients (r) between the "raw" pK_a's of the carboxylic acids listed in the third column (under the heading A) of Table I and the helical potentials of the 19 a-amino acids from Tanaka and Scheraga (1976); Chou and Fasman (1978); and Garnier, et al. (1978). The correlation coefficients are close to +0.60. The correlation was considerably improved by excluding all five charged residues (asp, glu, his, lys, arg) (Column b) or by merely excluding only two (asp and his) (Column c).

This improvement was not unexpected; earlier investigators had long ago recognized the helix-disrupting effects of charged groups carried by asp, his, and glu (Guzzo, 1965; Schiffer & Edmundson, 1967). I then asked the question, "Do these and other charged amino acid side chains all exist in an ionized, fully charged state (as for example, in a dilute solution of Na trichloracetate)?" An affirmative answer was presumed in the collection of data given in Column A of Table I. There is gathering evidence that this assumption may not be entirely correct:

(1) Kern (1939, 1948) showed that virtually all Na' in a dilute solution of Na isobutyrate is free; yet when these monomeric acid molecules are joined into a linear polymer, polyacrylate, the Na' in the solution'. becomes bound. Ling and Zhang (1983) have confirmed Kern's finding by demonstrating a similar and equally dramatic enhancement of Na' binding when (monomeric) styrene sulfonate is joined into the linear polymer, polystyrene sulfonate. These findings show that we cannot predict the degree of counterion association in polymers by examining the association with monomeric units (Ling, 1962, Chap. 1).

(2) Steinhardt & Zaiser (1951) showed that 36 β- and γ-carboxyl groups of ferri- and carbonyl-hemoglobin, which are not normally titratable in the native proteins, become titratable after acid denaturation. Very recently Ling and Zhang (1984) showed that in titration of bovine hemoglobin with NaOH, for each cationic group neutralized, one Na^+ becomes adsorbed onto the corresponding β- or γ-carboxyl groups thus "unmasked". This and other findings led to the conclusion that

in the native bovine oxy-hemoglobin and methemoglobin virtually all the anionic β- and γ-carboxyl groups are locked in salt linkages with the positively charged α-amino, ε -amino, and guanidyl groups. A similar, though less sweeping, concept was suggested long ago by Edsall & Wyman (1958). Demonstration of extensive salt linkage formation in deoxyhemoglobin and other proteins have been revealed by x-ray crystallography (Perutz, 1970).

These findings showed that many charged side chains in native proteins are associated with groups carrying the opposite electric charges and that the "raw" pK_a values of the five charged amino acids (asp, glu, his, lys, and arg) cited in Column A of Table I are in need of refinement. To achieve this I collected two sets of pK_a values for the α-carboxylic acid analogues of these five amino acids: one set corresponds to those of fully charged and ionized side chains; the other set corresponds to those with fully neutralized side chains. An average was then taken and shown in Column B of Table I and represented as $p\overline{K}_a$'s. The justifications for this seemingly arbitrary decision to use averages are threefold: (1) for reasons just given it is less arbitrary than using the raw pK_a's; (2) the charged groups engaged in a salt linkage ($-COO^- \ldots H_3N^+-$) are neither totally neutral as in the undissociated COOH (or NH_2) nor totally free as in $-COO^-$ or ($-NH_3^+$) but somewhere in between; (3) some charged groups may be free and/or associated with free counterions.

A few words need to be said about the way the two sets of dissociation constants were obtained for the five carboxylic acid analogues of asp, glu, his, arg, and lys. For asp and glu, there was little problem since the first and second dissociation constants of both succinic acid and malonic acid are known and they correspond respectively to the "neutral" and ionized pK_a's. It was however, necessary to add the value of 0.6 to the first dissociation constant due to the entropy factor ($RT \ln 2 = 0.6$) (Bjerrum, 1923). It was more difficult (if not impossible) to obtain experimentally the acid dissociation constants of the carboxyl groups of guanidylbutyric acid, δ-amino-n-valeric acid and imidazole-acetic acid with the cationic group in an *unionized* state: at the pH where the carboxyl groups titrate, the cationic groups are inevitably ionized. To circumvent this difficulty, I used Chiang and Tai's excellent method for calculating the pK_a's (and other functions) of organic molecules of known molecular structures (Chiang and Tai, 1962). Chiang and Tai's theory, not as well known in the West, permits accurate predictions of physicochemical properties as well as equilibrium and rate constants of organic compounds, based on the inductive concepts, molecular structure, and radii and electronegativity of the atoms involved (for a brief review of this theory see Chiang and Tai, 1985).

The last column (d) of Table II shows the correlation coefficients of the three sets of helical potential values against electron-donating power based on the improved dissociation constants of the carboxylic acid analogues, represented as $p\overline{K}_a$'s. The correlation coefficients of the electron-donating strength and the α-helical potentials are now +0.75 (Tanaka & Scheraga, 1976); 4-0.77 (Chou & Fasman, 1978); and +0.72 (Garnier et al., 1978), averaging +0.75. Considering that the accuracy of predicting new protein structure based on these and other lists of helical potentials, etc., are themselves in the range of 60 to 80%, the average correlation correction of +0.75 must be considered satisfactory. It is also worth noticing the closeness of the respective r's from the three sets of helical potentials, even though some differences exist among these sets of constants themselves; the correlation coefficients between each pair of helical potential constants for example, varied between +0.96 (Chou & Fasman vs. Garnier, et al) and +0.85 (Garnier, et al. vs. Tanaka & Scheraga).

Another corollary comment worth mentioning is that by evoking the much greater electrondonating strength of CH3 group than H, one can offer a simple explanation why glycine is a strong helix-breaker while alanine is a strong helix-former. Poly-L-alanine, which is insoluble in water, was brought into solution by being "sandwiched" between segments of soluble poly-D-glutamic and of poly-L-glutamic acid. In this block copolymer, the poly-L-alanine remained in the α-helical state even in the presence of 10 M urea. Yet as pointed out by Doty & Gratzer (1962), the alanine side chain is too short to form hydrophobic bonds with neighboring alanine side chains to enhance helical stability.

Next I examined the relation between the electron-donating power of the amino acid side chains represented by the $p\overline{K}_a$'s (and pK_a's) and the β-potentials (and the potentials for turns and coil). (All data derived from pK_a's are in parenthesis.) Column a of Table III shows that there is hardly any significant positive correlation between the @-potentials and the $p\overline{K}_a$'s, in apparent contradiction of Prediction 2 earlier cited. However, by excluding all five charged side chains (b) or only lys and glu (d) or glu alone (e) one obtained significantly improved correlation. Exclusion of his and asp (c) which strongly improved correlation between electron-donating power of the side chain and the α-helical potentials, on the other hand, has no effect on the correlation between electron-donating power of the side chain and the β-potentials. That the exclusion of the charged groups and in particular glu improved the correlation between $p\overline{K}_a$'s and @-potentialsis also reasonable since charged groups in general and glu in particular are long known to disrupt β-structure (Ptitsyn & Finkelstein, 1970; Chou & Fasman, 1974). The positive but weaker correlation between $p\overline{K}_a$'s and β-potentials than those between $p\overline{K}_a$'s and helical potentials are in harmony with predictions (Prediction 2). Similar though weaker correlations were demonstrated in all cases

		a	b	c	d	e
β leated sha	Tanaka and Scheraga	+0.37 (+0.19)	+0.68 (+0.68)	+0.38 (+0.38)	+0.55 (+0.30)	+0.49 (+0.30)
	Chou and Fasman	+0.10 (+0.07)	+0.55 (+0.55)	+0.10 (+0.07)	+0.42 (+0.15)	+0.34 (+0.16)
	Garnier, Osguthorpe and Robson	+0.16 (0.00)	+0.43 (+0.43)	+0.14 (+0.11)	+0.39 (+0.16)	+0.33 (+0.17)
Tur	Garnier, Osguthorpe and Robson	-0.62 (-0.45)	-0.74 (-0.74)	-0.70 (-0.73)	-0.60 (-0.38)	-0.57 (-0.38)
Coils	Garnier, Osguthorpe and Robson	-0.59 (-0.49)	-0.55 (-0.55)	-0.59 (-0.59)	-0.55 (-0.43)	-0.55 (-0.42)

TABLE III.[a] ***Linear correlation coefficients between acid dissociation constants of a-carboxylic acid analogues ($p\overline{K}_a$) and the potentials for extended conformation, coil, and turns.*** **Similar correlation coefficients between pK_a's and the various potentials are in parentheses. a: $p\overline{K}_a$'s and pK_a's here and elsewhere in this table are those given under Column B of Table I respectively. b: $p\overline{K}_a$'s and pK_a's with all five charged amino acid residues (asp, glu, lys, arg, and his) excluded. c: $p\overline{K}_a$'s and pK_a's with asp and his excluded. d: $p\overline{K}_a$'s and pK_a's with glu and lys excluded. e: $p\overline{K}_a$'s and pK_a's with glu (alone) excluded. $v_{c,j}$* from Tanaka and Scheraga (1976); $P\beta$ from 29 proteins from Chou and Fasman (1978) and directional information measure for extended conformation at residue position j (Garnier, et al., 1978).**

when pK_a's were used instead of $p\bar{K}_a$'s, in agreement with earlier conclusions from comparisons of correlations with a-helical potentials (Table II).

Significant negative correlations were also obtained between the electron-donating power of the side chains and the potentials for turns and coils given by Garnier et al. (1978) also in agreement with predictions (Prediction 3). Exclusions of charged groups here (b to f) have mild to no effect on these negative correlations. A reciprocal relation between helix potentials and coil potentials had already been noted and discussed by Robson & Suzuki (1976).

Side chains may also stabilize a-helical or β-structures by forming hydrophobic bonds with other hydrophobic side chains (Scheraga, 1961; Nemethy & Scheraga, 1962) or with other hydrophobic exteriors of other helices and extended structures in the formation of tertiary structures (Levitt & Chothia, 1976; Taylor & Thornton, 1983; Welinder et al., 1984). Therefore, I also determined the correlation between the index of hydrophobicity given by Tanford (1962) and the index of bulkiness by Zimmerman et al. (1968) with the potentials for helices, β-structures, turns and coils. The results are shown in Table IV. Note that there is good

		Hydrophobicity	Bulkiness
α-helix	Tanaka and Scheraga	+0.19	+0.46
	Chou and Fasman	+0.15	+0.35
	Garnier, Osguthorpe and Robson	+0.19	+0.40
β-pleated Sheet	Tanaka and Scheraga	+0.49	+0.63
	Chou and Fasman	+0.62	+0.76
	Garnier, Osguthorpe and Robson	+0.69	+0.67
Coil	Garnier, Osguthorpe and Robson	-0.60	-0.54
Turns	Garnier, Osguthorpe and Robson	-0.44	-0.51
	$p\bar{K}_a$	+0.27	+0.50

TABLE IV.[a] ***Linear correlation coefficients between index of bulkiness and of hydrophobicity vs. the potentials for a-helical conformation, extended conformation, coil and turns, and between indexes of bulkiness, hydrophobicity vs. $p\bar{K}_a$'s.*** **Index of bulkiness was taken from Zimmerman et al. (1968); index of hydrophobicity from Tanford (1962). Potentials for a-helical conformation are same as described under Table II. Potentials for coil and turns are the directional information measure for these structures at residue position j (Garnier, et al., 1978). $p\bar{K}_a$'s are those given under Column B of Table I.**

positive correlation of between 4-0.49 to 4-0.76 between hydrophobicity as well as bulkiness of the side chains and the β-structure potentials. Again this is reasonable since it has long been known that β-structures contain many hydrophobic residues (Ptitsyn & Finkelstein, 1970; Chou & Fasman, 1974, 1978). On the other hand, there is only weak positive correlation between hydrophobicity and a-helix potentials. In Table III, I have also shown that similar correlations exist between $\overline{pK}_a$'s and bulkiness as well as between $\overline{pK}_a$'s and hydrophobicity not unlike those shown between a-helix potentials and the hydrophobicity and bulkiness.

Thus as judged by the differences in the correlation coefficients it would seem that electron-donating power of the side chains plays a more dominant role than hydrophobicity or bulkiness in determining the a-helix potentials of the different side chains; while in the β-potential, the converse may be the case.

Both electron-donating power and hydrophobicity (and bulkiness) are also negatively correlated with the potentials for coils and turns and to about the same degree.

Now let us turn to another important question, whether or not the electron-donating power reaches beyond the amide groups of the same residue to other neighboring amide groups affecting their conformation (Postulate 4). For this investigation, I shall rely entirely on the work from Robson's laboratory in general and that of Garnier, Osguthorpe, and Robson (1978) in particular. The results are presented in Figure 1, in which linear correlation coefficients between $\overline{pK}_a$ of the jth residue and the a-helix potential of the jth site as well as the neighboring j+m sites are plotted. In the same figure, the same a-helical potentials are also plotted against

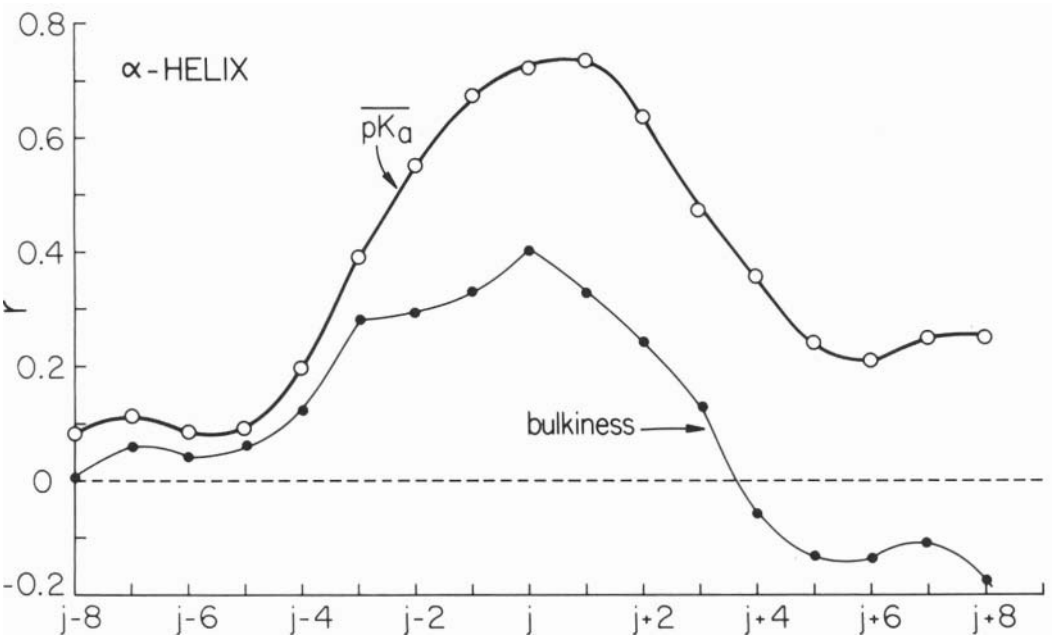


FIGURE I. Linear correlation coefficients (r) between a-helical potentials of 19 a-amino acids on the one hand; and on the other hand, the bulkiness as well as the electron-donating power of their separate side chains as expressed as the acid dissociation constants $\overline{pK}_a$'s of their a-carboxylic acid analogues at residue position j and j+m. m varied between –8 to +8. j–m represents a residue on the N-terminal side of the jth residue; j+m represents a residue on the C-terminal side of the jth residue. a-helical potentials are the directional information measure given by Garnier et al. (1978).

the bulkiness of the jth and neighboring side chain. Note the different contours of the two curves. The more or less smooth bell-shaped contours of the $p\overline{K}_a$ and its consistently positive r values are in harmony with the view that the electron-donating power of the jth residues spread out inductively along the polypeptide chains toward both the N and the C-terminals with diminishing effectiveness. In contrast, the bulkiness vs. the α-helical potentials plot shows more or less distinct peaks at j=0, j−3. and j+3, suggesting intra-helical hydrophobic bonds at these sites (Schiffer & Edmundson, 1967).

Since the bulkiness index is moderately correlated to $p\overline{K}_a$'s (Table 111), one raises the question, "How much of the observed correlation between the $p\overline{K}_a$ of the jth site and the a-helical potentials of the neighboring sites could reflect hydrophobic bond formation?" The answer depends on the site involved. At the jth site, the much higher r indicates that the dominating factor is the electron-donating strength. Since r^2 may be described approximately as the estimated proportion of the variance of helical potential that can be attributed to its linear regression on the electron-donating strength, half of the variance of the helical potential is associated with the electron-donating strength of the side chains of the jth residue. Using the same formula one finds that at the jth site, bulkiness accounts for only about 15% of the correlation.

At residues further away from the jth site, the correlation between the two curves of Figure I becomes quite noticeable; but in these remote areas the effect of propagated inductive effect of the jth side chain may also be expected to dwindle to low levels. There, relatively speaking, the hydrophobicity of the jth residue exercises a more pronounced effect. The fact that at j−4 and beyond (toward the C-terminal) the correlation between helical potential and bulkiness became negative probably arose from the inclusion of directional information that came from the neighboring non-helical segments (Garnier et al., 1978) since the average helical segment length is around ten residues (Chou & Fasman, 1978; Taylor & Thornton, 1983).

DISCUSSION

It is by now well established that each protein has a specific amino acid sequence. If every amino acid residue also retains specific potentials for helical and alternative conformations as repeatedly demonstrated, the protein's conformation would be permanent and unchanging. Yet it is also well established that physiological activities of living cells depend upon reversible and rapid changes of the conformation of cell proteins in response to the binding of ligands or "cardinal adsorbents" (e.g., drugs, hormones, ATP, Ca) on "receptor sites" of the proteins involved (Klotz, 1973; Aizono et al., 1974; Imae et al., 1975; Changeux, 1981). How can we then reconcile these two sets of apparently contradictory established facts? A key to the solution may again lie in the inductive effect, i.e., both the side chains and ligands bound to functional groups on side chains (and/or backbones) affect protein conformation by means of a similar inductive mechanism.

However, the secondary structures determined by the electron-donating strengths of side chains are primarily short-range; ligand induced changes of protein conformation are as a rule "global" and far-reaching (allosteric) resembling cooperative transitions during denaturation. Can the local short-range interaction give rise to far-reaching global effect by eliciting a cooperative transition? Some evidence exists suggesting that this is the case. Thus, the faulty substitution of one glu in hemoglobin A by val on the β-chains in hemoglobin S (Hunt & Ingram, 1959) produces profound sol-gel changes of the protein conformation (Itano, 1953). Chou (1974) discussed how an a-helical to β-structure alteration involving many amino acid residues in

hemoglobin might be predicted from the helical potentials and β-potentials of the two amino acid residues (glu and val) involved.

Garnier, Osguthorpe, and Robson (1978) also pointed out that the directional information they uncovered and on which the data of Figure I were based, contained in it the information of cooperativity among neighboring amino acid residues. The positive correlation demonstrated here between the directional information and the electron-donating strengths of the amino acid residues suggests the presence of an inductive component in the nearest neighbor interaction energy ($-\gamma/2$) in the helix-random coil transition [exp $(\gamma/RT) = 6$, where δ is Zimm and Bragg's initiative factor (Zimm & Bragg, 1958; Ling, 1964; 1980)]. The control of oxygen binding on hemoglobin by 2,3-diphosphoglycerate, inositol hexaphosphate, and ATP are well known examples of allosteric control by the ligands mediated by cooperative transitions (Chanutin & Curnish, 1964, 1967; Ling, 1970).

Witness also the improved correlation coefficient between $p\overline{K}_a$ values and α-helical potentials when the charged amino acid residues, his and asp, are partly neutralized (compare Table IIa and d). Such an improved correlation indicates that the partial neutralization of the charged group, for example, by ion binding may influence the conformation of the protein segment not unlike that brought on by the glu $-$ val substitution in hemoglobin S. The amino acid residues, asp and his. are of interest for another reason. They both possess short side chains comprising only one methylene group (which effectively reduces inductive effect transmission) separating their functional groups from the polypeptide chains. Thus the inductive effect originating in the changes of their functional groups may be expected to be partially transmitted to the nearby region of the protein molecules. Consider the Bohr effects which occur in hemoglobin as well as in ***single chain*** heme-proteins (Wald & Riggs, 1951). These effects demonstrate strong mutual dependence of the proton dissociation of some acidic groups of hemoglobin and oxygen binding on hemes attached to the histidine residues. There is evidence that at least in the reverse Bohr effect in mammalian hemoglobin, the interacting acidic groups are carboxyl (Rossi-Fanelli, et al.,1964). It is interesting that Perutz, whose elucidation of hemolobin structure galvanized the world, had suggested "in principle" a role of inductive effect in the control of oxygen binding on heme sites by the globin moiety of the hemoglobin molecule (Perutz, 1979). [A preliminary note has been published (Ling, 1984).]

The foregoing work was supported by NIH Grants 2-R01-CA16301 and 2-R01-GM11422 and Office of Naval Research Contract N00014-79-C-0126.

I thank Professor V. I. Lim, Inst. of Protein Research, Academy of Sciences of the USSR; Professor G. D. Fasman, Rosenfeld Professor of Biochemistry, Brandeis Univ.; Professor J. Garnier, Lab. de Biochimie Physique; Univ. de Paris-SUD; and Professor B. Robson, Dept. of Biochemistry, Univ. of Newcastle upon Tyne for reading the manuscript and for their valuable criticisms and suggestions.

REFERENCES

Aizono, Y., Roberts, J. E., Sonenberg, M., & Swislocki, N. I. (1974) *Arch. Biochem. Biophys.* 163, 634-643.

Anfinsen, C. B. (1962) *Brookhaven Symp. Biol.* 15, 184-194.

Anfinsen, C. B. (1967) *Harvey Lect.* 61, 95-116.

Bjerrum, N. (1923) *Z. physik. Chem.* 106, 219-224.

Branch, G. E. K., & Calvin, M. (1941) *The Theory of Organic Chemistry, An Advanced Course,* Prentice-Hall, Englewood Clifts.

Cannon, G. C. (1955) *Mikrochim Acta* 2-3, 555-560.

Changeux, J. P. (1981) *Harvey Lect.* 75, 85-254.

Chanutin, A., & Curnish, R. R. (1964) *Arch. Biochem. Biophys.* 106,433-439.

Chanutin, A., & Curnish, R. A. (1967) *Arch. Biochem. Biophys.* 121, 96-102.

Chapman, N. B., & Shorter, J. (1972) *Advances in Linear Free Energy Relationships,* Plenum Press, New York.

Chiang, M. C., & Tai, T. C. (1963) *Sci. Sin.* 12,785-866.
Chiang, M. C., & Tai, T. C. (1985) *Physiol. Chem. Phys. & Med. NMR* 17,271.
Chou, P. Y. (1974) *Biochem. Biophys. Res. Comm.* 61, 87-94.
Chou, P. Y., & Fasman, G. D. (1974) *Biochemistry* 13, 222-245.
Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol.* 47, 45-148.
Czarnetzky, E. J., & Schmidt, C. L. A. (1931) *J. Biol. Chem.* 92,453-454.
Davies, D. R. (1964) *J. Mol. Biol.* 9, 605-609.
Derick, C. G. (1911) J. *Amer. Chem. Soc.* 33, 1181-1189.
Dewar, M. J. S. (1949) *The Electronic Theory of Organic Chemistry,* Oxford University Press, London.
Doty, P., & Gratzer, W. B. (1962) in *Polyamino Acids, Polypeptides and Proteins* (Stahmann, M. A., Ed.) pp. 111-118, University of Wisconsin Press, Madison.
Edmundson, A. B. (1965) *Nature* 205, 883-887.
Edsall, J. T., & Wyman, J. (1958) *Biophysical Chemistry,* Vol. I, p. 539, Academic Press, New York.
Finkelstein, A. V., & Ptitsyn, O. B. (1971) J. *Mol. Biol.* 62, 613-624.
Garnier, J., Osguthorpe, D. J., & Robson, B. (1978) *J. Mol. Biol.* 120, 97-120.
Goldsack, D. E. (1969) *Biopolymers* 7,299-313.
Greenstein, J. P., & Winitz, M. (1961) in *Chemistry of the Amino Acids,* p. 486, Wiley & Sons, New York.
Guzzo, A. V. (1965) *Biophys. J.* 5,809-822.
Hammett, L. P. (1935) *Chem. Rev.* 17, 125-136.
Hammett, L. P. (1970) *Physical Organic Chemistry,* 2nd ed., McGraw-Hill, New York.
Hodgman, C. D., Weast, R. C., & Selby, S. M. (1961) *Handbook of Chemistry and Physics.* 1953-1956, The Chemical Rubber Publishing Company, Cleveland.
Hunt, J. A., & Ingram, V. M. (1959) *Nature* 184, 640-641.
Imae, T., Fasman, G. D., Hinkle, P. M., & Tashjian, A. H. (1975) *Biochem. Biophys. Res. Comm.* 62, 923-932.
Ingold, C. K. (1953) *Structure and Mechanism in Organic Chemistry,* Cornell University Press, Ithaca.
Itano, H. A. (1953) *Science* 117, 89-94.
Jencks, W. P., & Kegenstein, J. (1970) in *Handbook of Biochemistry,* 2nd ed. (Sober, H. A., Harte, R. A., & Sober, E. K., Eds) p. J191, The Chemical Rubber Publication Company, Cleveland.
Kauzmann, W. (1954) in *The Mechanism of Enzyme Action* (McElroy, W. D., and Glass, B., Eds.) pp. 70-110, Johns Hopkins University Press, Baltimore.
Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C., & Shore, V. C. (1960) *Nature* 185, 422-427.
Kern, W. (1939) Z. *phys. Chem.* Abt. A184, 197-210.
Kern, W. (1948) *Makramol. Chemie* (Akad. Verlag, Leipzig) 2,279-288.
Klotz, I. M. (1973) *Ann. N. Y. Acad. Sci.* 226, 18-35.
Kortum, G., Vogel, W., & Andrussow, K. (1961) *Pure Applied Chem.* 1, 187-536.
Kotelchuck, E., & Scheraga, H. A. (1969) *Proc. Natl. Acad. Sci.* 62, 14-21.
Levitt, M., & Chothia, C. (1976) *Nature* 261, 552-558.
Lewis, G. N. (1923) *Valence and Structure of Atoms and Molecules,* Chemical Catalogue, New York.
Lewis, P. N., & Scheraga, H. A. (1971) *Arch. Biochem. Biophys.* 144, 576-583.
Lim, V. I. (1974) J. *Mol. Biol.* 88, 857-871.
Ling, G. N. (1962) *A Physical Theory of the Living State — The Association-induction Hypothesis,* Blaisdell, Waltham.
Ling, G. N. (1964) *J. of Biopolymers* 1, 91-116.
Ling, G. N. (1970) *Proc. Natl. Acad. Sci.* 67, 296-301.
Ling, G. N. (1980) in *Cooperative Phenomenon in Biology* (Karreman, G., Ed.) pp. 39-69, Pergamon Press, New York.
Ling, G. N. (1984a) *In Search of the Physical Basis of Life,* Plenum Press, New York.
Ling, G. N. (1984b) *Physiol. Chem. Phys. & Med. NMR* 16,459-460.
Ling, G. N., & Zhang, Z. L. (1983) *Physiol. Chem. Phys. & Med. NMR* 15, 251-258.
Ling, G. N., & Zhang, Z. L. (1984) *Physiol. Chem. Phys. & Med. NMR* 16,221-235.
Low, B. W., Lovell, F. M., & Rudko, A. D. (1968) *Proc. Natl. Acad. Sci.* 60, 1519-1526.
Mizushima, S. (1954) *Structure of Molecules and Internal Rotation,* Academic Press, New York.
Nagano, K. (1974) J. *Mol. Biol.* 84,337-372
Nemethy, G., & Scheraga, H. A. (1962) J. *Phys. Chem.* 66, 1773-1789.
Pauling, L. (1960) *The Nature of the Chemical Bond,* 3rd ed. Cornell University Press, Ithaca.
Periti, P. F., Quagliarotti, G., & Liquori, A. M. (1967) *J. Mol. Biol.* 24. 313-322.
Perutz, M. F. (1970) *Nature* 228,726-739.
Perutz, M. F. (1979) *Ann. Rev. Biochem.* 48, 327-386.
Ponnuswarny, P. K., Warme, P. K., & Scheraga, H. A. (1973) *Proc. Natl. Acad. Sci.* 70, 830-833.
Prothero, J. W. (1966) *Biophys. J.* 6, 367-370.
Ptitsyn, O. B., & Finkelstein, A. V. (1970) *Biofizika* (U.S.S.R.) 15, 757-767.
Reiser, A. (1959) in *Hydrogen Bonding* (Hadzi, D., & Thompson, H. W., Eds.) Pergamon Press, New York.

Robson, B., & Pain, R. H. (1972) ***Nature*** 238, 107-108.

Robson, B., & Pain, R. H. (1974) ***Biochem. J.*** 141, 869-882.

Robson, B., & Suzuki, E. (1976) ***J. Mol. Biol.*** 107, 327-356.

Rogers, N. K., & Sternberg, M. E. (1984) ***J. Mol. Biol.*** 174, 527-542.

Rossi-Fanelli, A., Antonini, E., & Caputo, A. (1964) ***Adv. Protein Chem.*** 19, 73-222.

Scheraga, H. A. (1961) ***J. Phys. Chem.*** 65, 1071-1072.

Scheraga, H. A. (1974) ***Curr. Topics Biochem.*** 1-8.

Schiffer, M., & Edmundson, A. B. (1967) ***Biophys. J.*** 7, 121-135.

Sheridan, R. P., Levy, R. M., & Salemme, F. R. (1982) ***Proc. Natl. Acad. Sci.*** 79, 4545-4549.

Smith, R. P., Rhee, T., Magee, J. L., & Eyring, H. (1951) ***J. Amer. Chem. Soc.*** 73, 2263-3368.

Snedecor, G. W., & Cochran, W. G. (1980) ***Statistical Methods,*** 7th ed., Iowa State University Press, Ames.

Stecher, P. (1968) ***The Merck Index,*** 8th ed., Merck & Company, Inc., Rahway.

Steinhardt, J., & Zaiser, E. M. (1951) ***J. Biol. Chem.*** 190, 197-210.

Stiasny, E., & Scotti, H. (1930) ***Ber deut. chem. Ges.*** 63, 2977-2983.

Szent Györgyi, A. G., & Cohen, C. (1957) ***Science*** 126, 697-698.

Taft, R. W. (1960) ***J. Phys. Chem.*** 64, 1805-1815.

Taft, R. W., & Lewis, I. C. (1958) ***J. Amer. Chem. Soc.*** 80, 2436-2443.

Tanaka, S., & Scheraga, H. A. (1976) ***Macromol.*** 9, 168-182.

Tanford, C. (1962) ***J. Amer. Chem. Soc.*** 84,4240.

Taylor, W. R., & Thornton, J. M. (1983) ***Nature*** 301, 540-542.

Wada, A. (1976) ***Adv. Biophys.*** 9, 1-63.

Wald, G., & Riggs, A. (1951) ***J. Gen. Physiol.*** 35, 45-53.

Welinder, K. G., Mikkelsen, L., & Sottrup-Jensen, L. (1984) ***J. Biol. Chem.*** 259,8328-8331.

Wu, T. T., & Kabat, E. A. (1971) ***Proc. Natl. Acad. Sci.*** 68, 1501-1506.

Zimm, B. H., & Bragg, J. K. (1958) ***J. Chem. Phys.*** 28, 1246-1247.

Zimmerman, J. M., Eliezer, N., & Simha, R. (1968) ***J. Ther. Biol.*** 21, 170-201.

(Received March 6, 1986)